JIRSS (2015)
Vol. 14, No. 1, pp 1-23

Test of the Correlation Coefficient in Bivariate Normal Populations Using Ranked Set Sampling

Nader Nematollahi[1], **Reza Shahi**[2]

[1]Department of Statistics, Allameh Tabataba'i University, Tehran, Iran.
[2]Department of Mathematics, Tabriz Branch, Islamic Azad University, Iran.

Abstract. Ranked Set Sampling (RSS) is a statistical method for data collection that leads to more efficient estimators than competitors based on Simple Random Sampling (SRS). We consider testing the correlation coefficient of bivariate normal distribution based on Bivariate RSS (BVRSS). Under one-sided and two-sided alternatives, we show that the new tests based on BVRSS are more powerful than the usual uniformly most powerful tests based on bivariate SRS. Furthermore, the proposed tests for repeated and unrepeated samples are compared.

Keywords. Bivariate normal distribution, bivariate ranked set sampling, correlation coefficient, hypothesis testing, ranked set sampling.

MSC: 62F03, 62D05.

1 Introduction

In a real life situations when the measurements on the variable of interest is costly or time consuming and ranking of the sample items based on a correlated variable can be easily done, by judgment without actual measurement, RSS method can be used and it is superior to the SRS method. The RSS was first proposed by McIntyer (1952) for estimating the mean of pasture yields. Takahasi and Wakimoto (1968) described

Nader Nematollahi(✉)(nematollahi@ata.ac.ir), Reza Shahi(shahi@iaut.ac.ir)
Received: April 2014; Accepted: April 2015

the mathematical theory of RSS. Dell and Clutter (1972) showed that the RSS mean is unbiased and more efficient than the SRS mean even when the ranking is imperfect. RSS strategy has emerged as a powerful tool and a serious alternative to the commonly used SRS methods.

In the past two decades, many researchers used the RSS method in both parametric and nonparametric problems and extended some modification for RSS. For example, in parametric inference Ni Chuiv et al. (1994), Fei et al. (1994), Lam et al. (1994, 1995), Sinha and Purkayastha (1996) and Bhoj and Ahsanullah (1996) used RSS to estimate the parameters of Cauchy, Weibull, extreme-value, two parameter exponential, logistic, normal, exponential and generalized geometric distributions, respectively. In modification of RSS, Samawi et al. (1996), Muttlak (1997), Li et al. (1999), Hossain and Muttlak (2001) and Rahimov and Muttlak (2003) suggested extreme RSS, median RSS, random selection in RSS, selected RSS (SRSS) and generalization of random selection in RSS, respectively. Chen et al. (2004) provides a detailed discussion of RSS and its variants.

For hypothesis testing of the parameter of some distributions based on one characteristic, some authors used RSS to obtain the tests, which were more powerful than the usual Uniformly Most Powerful (UMP) tests based on SRS. Among them, Abu-Dayyeh and Muttlak (1996), Muttlak and Abu-Dayyeh (1998), Wang and Tseng (2002), Tseng and Wu (2007), Hossain and Muttlak (2006) and Hossain and Khan (2006) used RSS, median RSS and SRSS to test one of the parameters of normal, exponential, uniform and rectangular distributions.

A few researchers have considered estimating multiple characteristics using RSS. One way of using RSS for this case is ranking the units by judgment with respect to one chosen characteristic, and all other characteristics are given the same ranking order as the ranked characteristic. McIntyer (1952), Takahasi (1970), Stokes (1980), Patil et al. (1994), Zheng and Modarres (2006) and Hui et al. (2009) used this approach. Stokes (1980) applied mentioned method and showed that the maximum likelihood estimator of the correlation coefficient ρ of the Bivariate Normal (BVN) distribution based on RSS is asymptotically as efficient as the one obtained by SRS. Zheng and Modarres (2006) used this method to find a robust estimate of ρ of the BVN distribution. Another way for using RSS in multiple characteristics is to rank the units based on ordering all characteristics. Al-Saleh and Zheng (2002) introduced this method for ranking the units based on ordering of two characteristics and referred to it as Bivariate RSS (BVRSS). Al-Saleh and Samawi (2005) used BVRSS to estimate the correlation coefficient ρ of two characteristics and applied it to estimate ρ of the BVN distribution.

Suppose (X, Y) is a random vector which has a bivariate distribution with joint probability density function (p.d.f.) $f_{X,Y}(x, y)$ and ρ is the correlation coefficient between X and Y. The method of Al-Saleh and Zheng (2002) for obtaining a BVRSS of size k^2 is given by the following steps:

1. A random sample of size k^4 is chosen from the population and is randomly allocated into k^2 pool of size k^2, where each pool is a square matrix with k rows and k columns.

2. In the first pool, minimum value is chosen by judgment with respect to the first characteristic, for each of the rows.

3. For the k minima obtained in step 2, the pair that corresponds to the minimum value of the second characteristic is selected by judgment for actual quantification. This pair, which denoted by the label $(1, 1)$, is the first element of BVRSS.

4. Steps 2 and 3 is repeated for the second pool, but in step 3, the pair that corresponds to the second minimum value is selected for actual quantification with respect to the second characteristic. This pair is denoted by the label $(1, 2)$.

5. The process is continued until the label (k, k) resembles to the k^2th (last) pool.

This process produces a cycle of BVRSS of size k^2. If a sample of higher size is required, then this cycle could be repeated m times until the required size $n = mk^2$ is achieved.

It must be born in mind that in spite of the fact that k^4 units are identified for the BVRSS sample (for $m = 1$), only k^2 are selected for actual quantification. All k^4 units, however, contribute information to the k^2 quantified units. Al-Saleh and Zheng (2002) pointed out that when we have concomitant variables which are correlated to the variables (X, Y) and could be easily ranked without actual quantification, then we could use these variables in the ranking procedure instead of variables of interest.

To the best of our knowledge, the BVRSS for hypothesis testing on the parameters of bivariate distributions is not used in the literature. In the present paper, we construct some new test functions for testing correlation coefficient ρ of a bivariate normal distribution based on BVRSS, and show that these tests are more powerful than usual UMP Unbiased (UMPU) tests.

To this end, in Section 2 we consider the hypothesis testing on ρ based on BVRSS and compare the power function of these tests with respect to UMPU tests based on SRS. In Section 3 the tests based on repeated and unrepeated samples are compared. In Section 4, we use the asymptotic distribution of the sample correlation coefficient based on BVRSS to find the approximate value of cutoff points of the test statistic. In Section 5, using bootstrap method, the approximated p-value of the proposed test is computed based on BVRSS and is compared with p-value of the test based on BVSRS. A conclusion is given in Section 6.

2 Testing the Correlation Coefficient

In this section, based on a BVRSS of size $n = mk^2$ from a bivariate normal distribution with correlation coefficient ρ, we would like to test

$$H_0 : \rho = 0 \quad \text{versus} \quad H_0 : \rho \neq 0 \tag{1}$$

Suppose that $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ is a Bivariate Simple Random Sample (BVSRS) of size n from a population with $E(X_i) = \mu_x$, $E(Y_i) = \mu_y$, $V(X_i) = \sigma_x^2$, $V(Y_i) = \sigma_y^2$ and $\text{Corr}(X_i, Y_i) = \rho$, $i = 1, \cdots n$. Assume that μ_x, μ_y, σ_x^2, σ_y^2 and ρ are unknown. When the sample are drowned from BVN distribution $N_2(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$, it could be shown that the UMPU size α test for testing (1) is given by (Shao, 1999)

$$\varphi(r_{\text{BVSRS}}) = \varphi_1(r) = \begin{cases} 1 & \dfrac{|r|\sqrt{n-2}}{\sqrt{1-r^2}} > t_{1-\frac{\alpha}{2}}(n-2) \\ 0 & \text{o.w} \end{cases} \tag{2}$$

where

$$r_{\text{BVSRS}} = r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\left\{\left(\sum_{i=1}^{n}(X_i - \overline{X})^2\right)\left(\sum_{i=1}^{n}(Y_i - \overline{Y})^2\right)\right\}^{\frac{1}{2}}} \tag{3}$$

is an estimator of ρ and $t_{1-\frac{\alpha}{2}}(n-2)$ is the upper $\dfrac{\alpha}{2}$ percentile of the t distribution with $n-2$ degrees of freedom.

Now, to test the hypothesis (1) based on BVRSS, we first use BVRSS of size k^2 of Al-Saleh and Zheng (2002) which is introduced in Section 1. Suppose (X, Y) is a random vector which has BVN distribution $N_2(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$. A random sample of size k^4 is chosen from this population and is randomly allocated into k^2 square pools each have size k^2. It is assumed that the elements of each pool are randomly divided into k sets of size k. Denote the values of the two characteristics of the elements in the kth pool by $\{(X_{ij}^h, Y_{ij}^h), i = 1, \cdots, k,\ \ j = 1, \cdots, k,\ \ h = 1, 2, \cdots, k^2\}$, in which X_{ij}^h is the jth element of the ith row in the hth pool for the first characteristic and Y_{ij}^h is defined in the same manner for the second characteristic. Suppose that $X_{i(j)}^h$ is the jth minimum of the elements in the ith row in the hth pool, where $i = 1, 2, \cdots, k$, $j = 1, 2, \cdots, k$ and $h = (j-1)k + 1, \cdots, jk$, and $Y_{i[j]}^h$ is the corresponding value of Y. Finally, suppose that $Y_{(i)[j]}^h$ is the ith minimum of the elements $Y_{i[j]}^h$, $i = 1, 2, \cdots, k$ and $X_{[i](j)}^h$ is the corresponding value of X. Thus the BVRSS sample consists of k^2 pairs $(X_{[i](j)}^h, Y_{(i)[j]}^h)$ where $i = 1, 2, \cdots, k$, $j = 1, 2, \cdots, k$ and $h = (j-1)k + 1, \cdots, jk$, which are independent but not identically distributed. In the present paper, we assume that the judgment ranking is perfect; therefore, the ith judgment order statistic is the same as the ith order statistic. Note that the small brackets on subscripts are used to show that the ordering is perfect while the square brackets are used to indicate that the ordering is with respect to the perceived ranks induced by the other variable (concomitant variable). Obviously, the index h could be dropped.

Suppose that $(X_{[i](j)}, Y_{(i)[j]})$, $i = 1, 2, \cdots, k$, $j = 1, 2, \cdots, k$ is a BVRSS of size k^2 from BVN distribution. Similar to (3), Al-Saleh and Samawi (2005) proposed the following estimator of ρ based on BVRSS

$$r_{\mathrm{BVRSS}} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{k}(X_{[i](j)} - \overline{X}_{\mathrm{BVRSS}})(Y_{(i)[j]} - \overline{Y}_{\mathrm{BVRSS}})}{\left\{\left[\sum_{i=1}^{k}\sum_{j=1}^{k}(X_{[i](j)} - \overline{X}_{\mathrm{BVRSS}})^2\right]\left[\sum_{i=1}^{k}\sum_{j=1}^{k}(Y_{(i)[j]} - \overline{Y}_{\mathrm{BVRSS}})^2\right]\right\}^{\frac{1}{2}}} \tag{4}$$

where

$$\overline{X}_{\mathrm{BVRSS}} = \widehat{\mu}_{x\mathrm{BVRSS}} = \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k} X_{[i](j)} \quad ,$$

$$\overline{Y}_{\mathrm{BVRSS}} = \widehat{\mu}_{y\mathrm{BVRSS}} = \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k} Y_{(i)[j]}$$

Remark 2.1. There exist other estimators of ρ in the literature. Among them Stokes (1980) and Zheng and Modarres (2006) proposed two estimators of ρ based on RSS. In Appendix we compare these estimators with r_{BVSRS} in (3) and r_{BVRSS} in (4) by a simulation study. The results show that r_{BVRSS} in (4) is more efficient than the mentioned estimators of ρ. Note that the estimators of ρ introduced by Zheng and Modarres (2006) and Stokes (1980) are obtained by ranking a chosen characteristic and the other characteristic has the same ranking order as the ranked characteristic. These estimators are designed (by ignoring part of likelihood function) to be more robust than the BVSRS estimator. But the estimator r_{BVRSS} of Al-Saleh and Samawi (2005) is obtained by ranking the units based on two characteristics, and by a simulation study, we show that r_{BVRSS} is more efficient than the other estimators. Therefore, we use r_{BVRSS} for testing the hypothesis (1).

Remark 2.2. As noted in Section 1, the estimator (4) is obtained from one cycle ($m = 1$) of BVRSS. If we repeat this cycle m times, we will achieve a BVRSS of size $n = mk^2$. We use the term "unrepeated" and "repeated" sampling for the cases $m = 1$ and $m > 1$, respectively. In the case of repeated BVRSS, the estimator (4) changes to

$$r_{\text{BVRSS},m} = \frac{\sum_{s=1}^{m}\sum_{i=1}^{k}\sum_{j=1}^{k}(X_{[i](j)s} - \overline{X}_{\text{BVRSS},m})(Y_{(i)[j]s} - \overline{Y}_{\text{BVRSS},m})}{\left\{\left[\sum_{s=1}^{m}\sum_{i=1}^{k}\sum_{j=1}^{k}(X_{[i](j)s} - \overline{X}_{\text{BVRSS},m})^2\right]\left[\sum_{s=1}^{m}\sum_{i=1}^{k}\sum_{j=1}^{k}(Y_{(i)[j]s} - \overline{Y}_{\text{BVRSS},m})^2\right]\right\}^{\frac{1}{2}}} \quad (5)$$

where $(X_{[i](j)s}, Y_{(i)[j]s})$, $i = 1, 2, \cdots, k$, $j = 1, 2, \cdots, k$ is a BVRSS of size k^2 in the sth cycle, $s = 1, 2, \cdots, m$, and

$$\overline{X}_{\text{BVRSS},m} = \frac{1}{mk^2}\sum_{s=1}^{m}\sum_{i=1}^{k}\sum_{j=1}^{k} X_{[i](j)s} \quad ,$$

$$\overline{Y}_{\text{BVRSS},m} = \frac{1}{mk^2}\sum_{s=1}^{m}\sum_{i=1}^{k}\sum_{j=1}^{k} Y_{(i)[j]s} \ .$$

For testing (1) based on BVRSS, we propose the following test function

$$\varphi_2(r_{\text{BVRSS},m}) = \begin{cases} 1 & \dfrac{|r_{\text{BVRSS},m}|\sqrt{n-2}}{\sqrt{1 - r^2_{\text{BVRSS},m}}} > l_1(n - \alpha) \\ 0 & \text{o.w} \end{cases} , \quad (6)$$

where $n = mk^2$ and $l_1(n, \alpha)$ is determined by solving the equation $E_{\rho=0}[\varphi_2(r_{\text{BVRSS},m})] = \alpha$. The power function of $\varphi_2(r_{\text{BVRSS},m})$ is given by

$$E_\rho[\varphi_2(r_{\text{BVRSS},m})] = P\left(\frac{|r_{\text{BVRSS},m}|\sqrt{n-2}}{\sqrt{1-r^2_{\text{BVRSS},m}}} > l_1(n-\alpha)\right).$$

It could be observed from Figure 1 that this function is a decreasing (an increasing) function of ρ for $\rho \in [-1, 0]$ ($\rho \in [0, 1]$).

Since the value of $l_1(n, \alpha)$ could not be determined analytically, we compute the values of $l_1(n, \alpha)$ for $n = 4, 8, 9, 12, 16, 18, 20, 24, 25, 28, 32$ and $\alpha = 0.01, 0.05, 0.1$ by a simulation study. For each n, we draw a BVRSS of size n from $N_2(-1, 1, 1, 1, \rho = 0)$, $N_2(-1, 1, 2, 4, \rho = 0)$ and $N_2(1, 2, 1, 1, \rho = 0)$ distributions and compute $t_{\text{BVRSS}} = \dfrac{|r_{\text{BVRSS},m}|\sqrt{n-2}}{\sqrt{1-r^2_{\text{BVRSS},m}}}$. Then we repeat the sampling procedure 5000 times to determine $l_1(n, \alpha)$ from the $1 - \alpha$ percentile of the t_{BVRSS} values. The process is then repeated 2500 times and the average of those $l_1(n, \alpha)$'s is the desired values of $l_1(n, \alpha)$.

The simulated values of $l_1(n, \alpha)$ in test function (6) are computed for the mentioned BVN distributions with $\rho = 0$, and are denoted by $l_{1i}(n, \alpha)$, $i = 1, 2, 3$, respectively. These values are given in Table 1. The computation (and other computations in preceding sections) is done using R 2.15.3 software. Also, the values of $t_{1-\frac{\alpha}{2}}(n-2)$ are given in Table 1. Note that, for choosing a BVRSS of size $n = 8, 12, 18, 20, 28, 32$ we use repeated BVRSS method explained in Remark 2.2. For instance, when $n = 32$, we choose a sample of size $k^2 = 2^2$ and repeat the sampling $m = 8$ cycles, i.e., $n = mk^2 = 8(2^2) = 32$.
From Table 1, we observe that the values of $l_{1i}(n, \alpha)$, $i = 1, 2, 3$ are close to each other.

In order to compare the power of the test function (6) with power of (2) with $n = mk^2$, the simulation study is carried out based on the sample from BVN distribution. Let $\beta_1^*(\rho)$ and $\beta_2^*(\rho)$ be the power functions of $\varphi_1(r_{\text{BVSRS}})$ and $\varphi_2(r_{\text{BVRSS},m})$, respectively. We compute $\beta_1^*(\rho)$ and $\beta_2^*(\rho)$ for $\rho = -1$ to $\rho = 1$ at an increment 0.05 by a simulation study. For each n and ρ, we draw a BVRSS of size n from $N_2(-1, 1, 1, 1, \rho)$ and compute t_{BVRSS}. Then we repeat the sampling procedure 5000 times and determine the value of power function by the percent of t_{BVRSS}'s those that are greater than $l_1(n, \alpha)$. The process

Table 1: Critical values of the size-α BVSRS and BVRSS-based tests for a correlation coefficient of bivariate normal $N_2(-1,1,1,1,\rho=0)$, $N_2(-1,1,2,4,\rho=0)$ and $N_2(1,2,1,1,\rho=0)$

			$\alpha = 0.01$			
m	k	n	$t_{1-\frac{\alpha}{2}}(n-2)$	$l_{11}(n,\alpha)$	$l_{12}(n,\alpha)$	$l_{13}(n,\alpha)$
1	2	4	9.9248	7.6444	8.1152	8.4171
2	2	8	3.7074	3.0395	3.3464	3.5088
1	3	9	3.4995	2.7004	3.0598	3.1433
3	2	12	3.1693	2.6895	3.0382	3.0251
4	2	16	2.9768	2.0073	2.3974	2.4280
2	3	18	2.9208	2.3270	2.5001	2.5874
5	2	20	2.8784	2.5420	2.6702	2.7662
6	2	24	2.8188	2.4558	2.6749	2.6430
1	5	25	2.8073	1.7735	1.9997	2.1086
7	2	28	2.7787	2.4604	2.5321	2.5974
8	2	32	2.7500	2.4920	2.5488	2.5905

			$\alpha = 0.05$			
m	k	n	$t_{1-\frac{\alpha}{2}}(n-2)$	$l_{11}(n,\alpha)$	$l_{12}(n,\alpha)$	$l_{13}(n,\alpha)$
1	2	4	4.3026	3.8588	3.9160	3.9879
2	2	8	2.4469	2.1920	2.2074	2.2930
1	3	9	2.3646	1.8712	1.9746	2.1760
3	2	12	2.2281	2.0469	2.0894	2.1483
4	2	16	2.1448	1.6327	1.7933	1.8675
2	3	18	2.1199	1.7800	1.8001	1.9930
5	2	20	2.1009	1.9406	1.9611	1.9877
6	2	24	2.0739	1.8955	1.9259	1.9580
1	5	25	2.0687	1.4459	1.6745	1.8944
7	2	28	2.0555	1.8499	1.8992	1.9659
8	2	32	2.0423	1.8456	1.8916	1.9157

			$\alpha = 0.1$			
m	k	n	$t_{1-\frac{\alpha}{2}}(n-2)$	$l_{11}(n,\alpha)$	$l_{12}(n,\alpha)$	$l_{13}(n,\alpha)$
1	2	4	2.9200	2.5418	2.6327	2.7463
2	2	8	1.9432	1.7673	1.7764	1.8516
1	3	9	1.8946	1.6044	1.6407	1.7010
3	2	12	1.8125	1.6072	1.6719	1.7494
4	2	16	1.7613	1.5035	1.6139	1.6697
2	3	18	1.7459	1.4759	1.5064	1.6109
5	2	20	1.7341	1.5648	1.6034	1.6612
6	2	24	1.7171	1.5581	1.5874	1.6107
1	5	25	1.7139	1.3163	1.3940	1.5077
7	2	28	1.7056	1.5485	1.5844	1.6348
8	2	32	1.6973	1.5466	1.5559	1.5903

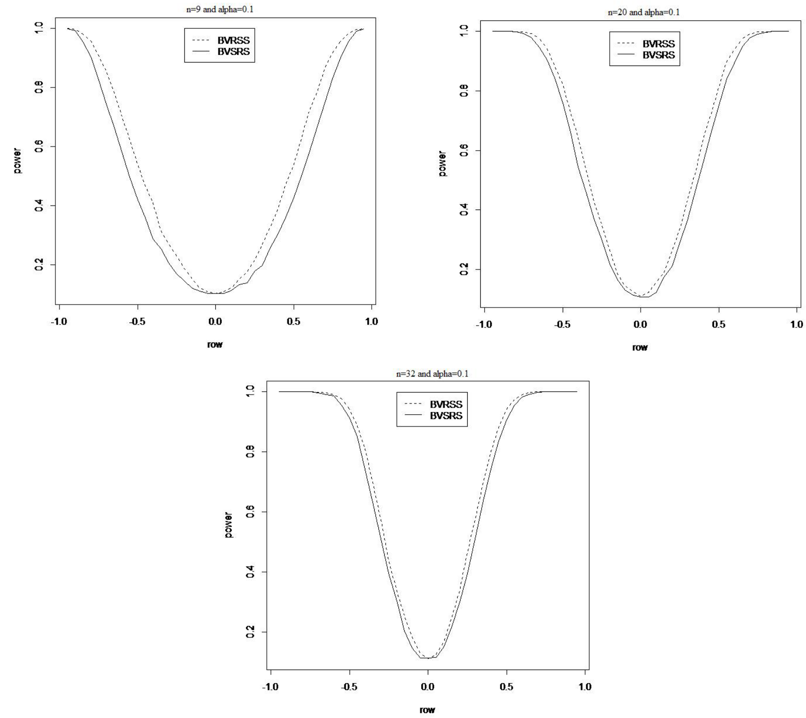

FIGURE 1. Plots of $\beta_1^*(\rho)$ and $\beta_2^*(\rho)$ for $\alpha = 0.1$ and $n = 9, 20, 32$

is then repeated 2500 times and the average of these percents is the desired value of the power function. The simulation is carried out for $n = 4, 8, 9, 12, 16, 18, 20, 24, 25, 28, 32$ and $\alpha = 0.01, 0.05, 0.1$. In Figure 1, the plot of simulated $\beta_1^*(\rho)$ and $\beta_2^*(\rho)$ are given for $n = 9, 20, 32$ and $\alpha = 0.1$. All other cases have similar figures and hence are omitted. From these plots we observe that $\varphi_2(r_{\text{BVRSS},m})$ is more powerful than the usual UMPU test $\varphi_1(r_{\text{BVSRS}})$.

In order to have a more quantitative comparison, we use the Maximum Rate of Improvement (MRI) criterion. For doing this let for a given n, $\rho_1^* = \rho^*(n)$ be any value that maximize $\beta_2^*(\rho) - \beta_1^*(\rho)$ with respect to ρ. Then the MRI is defined by

$$MRI = \frac{\beta_2^*(\rho_1^*) - \beta_1^*(\rho_1^*)}{\beta_1^*(\rho_1^*)} \times 100.$$

The comparison between $\varphi_2(r_{\text{BVRSS},m})$ and $\varphi_1(r_{\text{BVSRS}})$ in terms of MRI, for $n = 4, 8, 9, 12, 16, 18, 20, 24,\ 25, 28, 32$ is presented in Table 2. Note that all values of MRI are above 7.63%, which indicate that the improvement of $\varphi_2(r_{\text{BVRSS},m})$ over $\varphi_1(r_{\text{BVSRS}})$ could be substantial, hence practically significant.

Table 2: The comparison between $\beta_1^*(\rho_1^*)$ and $\beta_2^*(\rho_1^*)$ and the maximum rate of improvement ($\alpha = 0.1$)

m	k	n	ρ_1^*	$\beta_2^*(\rho_1^*)$	$\beta_1^*(\rho_1^*)$	MRI
1	2	4	0.80	0.5148	0.3747	37.28
2	2	8	-0.35	0.2921	0.2714	7.63
3	2	12	0.60	0.7110	0.6218	14.35
4	2	16	0.65	0.7383	0.6819	8.27
5	2	20	-0.40	0.5460	0.4438	23.03
6	2	24	0.30	0.6009	0.5276	13.89
7	2	28	0.45	0.8551	0.7811	9.47
8	2	32	0.25	0.6840	0.5946	15.04

For the one sided test $H_0 : \rho \leq 0$ versus $H_0 : \rho > 0$, we use the test statistic $\dfrac{|r_{\text{BVRSS},m}|\sqrt{n-2}}{\sqrt{1-r^2_{\text{BVRSS},m}}}$ and derive the similar results. The details are omitted. For a special case, see Section 3 below.

3 Comparing the Tests Based on Repeated and Unrepeated Samples

The success of BVRSS procedure depends on the ability of ranking the k units correctly. Therefore, the size k in most practical applications is selected no larger than five or six. As noted in Section 1, in order to increase effective sample sizes, we choose BVRSS of size k^2 with small value of k, and then repeat sampling m cycles to achieve $n = mk^2$ sample (see Remark 2.2). Now if we construct the tests based on repeated and unrepeated BVRSS, which of them are more powerful? To answer this question, let the sample size be $n = 16$. Then we could choose the sample in the following ways:

(I) Choose BVRSS of size $k^2 = 2^2$ and repeat it $m = 4$ cycles, i.e., $n = mk^2 = 4(2^2) = 16$. (repeated BVRSS)

(II) Choose BVRSS of size $n = k^2 = 4^2 = 16$. (unrepeated BVRSS, $m = 1$).

For comparing these two cases, we consider the following hypothesis testing

$$H_0 : \rho \leq 0 \quad \text{versus} \quad H_1 : \rho > 0. \tag{7}$$

Based on BVSRS $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ from $N_2(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$, a UMPU size α-test for testing (7) is given by (Shao, 1999)

$$\varphi_3(r_{\text{BVSRS},m}) = \varphi_3(r) = \begin{cases} 1 & \dfrac{r\sqrt{n-2}}{\sqrt{1-r}} > t_{1-\alpha}(n-2) \\ 0 & \text{o.w} \end{cases}, \tag{8}$$

where r is given in (3) with $n = mk^2$. Similar to the previous section, we propose the following test based on BVRSS from a BVN distribution for testing (7),

$$\varphi_4(r_{\text{BVRSS},m}) = \begin{cases} 1 & \dfrac{r_{\text{BVRSS},m}\sqrt{n-2}}{\sqrt{1-r^2_{\text{BVRSS},m}}} > d_1(n-\alpha) \\ 0 & \text{o.w} \end{cases}, \tag{9}$$

where $r_{\text{BVRSS},m}$ is given in (5) and $d_1(n, \alpha)$ is determined by solving the equation

$$E_{\rho=0}[\varphi_4(r_{\text{BVRSS},m})] = \alpha.$$

We construct a simulation study by choosing BVRSS in the above two cases and compute $d_1(n, \alpha)$ in (9) for each cases. Table 3 shows the values of $d_1(n, \alpha)$ for the cases I and II.

Table 3: Critical values of the size-α repeated and unrepeated BVRSS-based test for a correlation coefficient of bivariate normal distribution

	$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.1$	
n	repeated	unrepeated	repeated	unrepeated	repeated	unrepeated
16	1.8817	1.4936	1.5035	1.1875	1.1914	1.0051

Moreover, we plot the power functions of BVRSS tests, i.e., $\beta_4^*(\rho) = P_\rho\left(\dfrac{r_{\text{BVRSS},m}\sqrt{n-2}}{\sqrt{1-r^2_{\text{BVRSS},m}}} > d_1(n, \alpha)\right)$ in the cases I and II and the UMPU

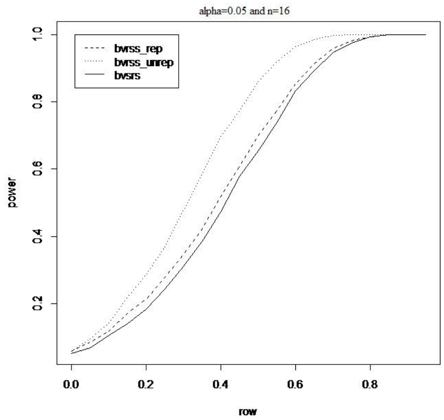

FIGURE 2. Plots of $\beta_3^*(\rho)$, $\beta_4^*(\rho)$ for repeated and unrepeated sampling with $\alpha = 0.05$ and $n = 16$.

test $\varphi_3(r_{\text{BVSRS}})$ in Figure 2. The plots "bvrss-rep", "bvrss-unrep" and "bvsrs" show the plot of power functions of $\varphi_4(r_{\text{BVRSS},m})$ in the cases I and II and $\varphi_3(r_{\text{BVSRS}})$, respectively. From Figure 2 we observe that the power function of $\varphi_4(r_{\text{BVRSS},m})$ in case II is higher than the power function of $\varphi_4(r_{\text{BVRSS},m})$ in case I, and both of them are higher than the power function of $\varphi_3(r_{\text{BVSRS}})$.

Although in the case of unrepeated sample, the test is more powerful, but with regard to error in ranking for large sample size, it is recommended to use the test with repeated sampling for large sample size. In this case the test is more powerful than UMPU test and only caused a reduction of power with respect to unrepeated sampling.

4 Simulation-Free Formula for Cutoff Points

In Sections 3 and 4, the values of the cutoff points of the proposed tests for testing correlation coefficient are evaluated using a simulation method for $n = 4, 8, 9, 12, 16, 18, 20, 24, 25, 28, 32$. When the sample size gets larger (for example, $n = 20, 24, 25, 28, 32, \cdots$), using simulation method for obtaining one value of the cutoff point needs a considerable amount of time. Regarding this difficulty, we use the asymptotic distribution of r_{BVRSS} to find a formula for computing the cutoff points for our proposed BVRSS-based tests. We do this for repeated BVRSS-based tests.

Suppose that (X_{ijs}, Y_{ijs}), $i = 1, 2, \cdots, k$, $j = 1, 2, \cdots, k$, $s = 1, 2, \cdots, m$ is a BVSRS of size $n = mk^2$ from the joint probability density function (p.d.f.) $f_{X,Y}(x, y)$. Let μ_x, μ_y, σ_x^2, σ_y^2 and ρ be the means and variances of X and Y, and the correlation coefficient between X and Y. We assume that μ_x, μ_y, σ_x^2, and σ_y^2 are known and ρ is unknown and we attempt to test the hypothesis (1). Without loss of generality we assume that $\mu_x = \mu_y = 0$ and $\sigma_x = \sigma_y = 1$. This implies that $\rho = \text{cov}(X, Y) = E(XY)$. An unbiased estimator of ρ based on this sample is

$$\widehat{\rho}_{\text{BVSRS},m} = \frac{\sum_{s=1}^{m}\sum_{i=1}^{k}\sum_{j=1}^{k}(X_{ijs}Y_{ijs})}{mk^2}. \tag{10}$$

Now suppose that $(X_{[i](j)s}, Y_{(i)[j]s})$, $i = 1, 2, \cdots, k$, $j = 1, 2, \cdots, k$, $s = 1, 2, \cdots, m$ is a repeated BVRSS of size $n = mk^2$ from the joint p.d.f. $f_{X,Y}(x, y)$ with $\mu_x = \mu_y = 0$ and $\sigma_x = \sigma_y = 1$. Similar to (10), Al-Saleh and Samawi (2005) proposed the following unbiased estimator of ρ based on BVRSS

$$\widehat{\rho}_{\text{BVRSS},m} = \frac{\sum_{s=1}^{m}\sum_{i=1}^{k}\sum_{j=1}^{k}(X_{[i](j)s}Y_{(i)[j]s})}{mk^2} = \frac{1}{m}\sum_{s=1}^{m}\widehat{\rho}_{\text{BVRSS},s} \tag{11}$$

where $\widehat{\rho}_{\text{BVRSS},s} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{k}(X_{[i](j)s}Y_{(i)[j]s})}{k^2}$, $s = 1, \cdots, m$. Note that $\widehat{\rho}_{\text{BVRSS},1}, \widehat{\rho}_{\text{BVRSS},2}, \cdots, \widehat{\rho}_{\text{BVRSS},m}$ are independent and identically distributed, with common mean ρ and variance $\text{var}(\widehat{\rho}_{\text{BVRSS}})$ which is given by formula (10) of Al-Saleh and Samawi (2005). Also Al-Saleh and Samawi (2005) demonstrated that for a fixed set size k,

$$\sqrt{m}(\widehat{\rho}_{\text{BVRSS},m} - \rho) \overset{in\ dist}{\longrightarrow} N(0, \text{var}(\widehat{\rho}_{\text{BVRSS}})).$$

Therefore,

$$\frac{\sqrt{m}(\widehat{\rho}_{\text{BVRSS},m} - \rho)}{\sqrt{\text{var}(\widehat{\rho}_{\text{BVRSS}})}} \overset{in\ dist}{\longrightarrow} N(0, 1). \tag{12}$$

For testing (1) based on repeated BVRSS, we propose the following test function

$$\varphi_5(\widehat{\rho}_{\mathrm{BVRSS},m}) = \begin{cases} 1 & \dfrac{|\widehat{\rho}_{\mathrm{BVRSS},m}|\sqrt{mk^2-2}}{\sqrt{1-\widehat{\rho}^2_{\mathrm{BVRSS},m}}} > c(n-\alpha) \\ 0 & \mathrm{o.w} \end{cases}, \tag{13}$$

where $c(n,\alpha)$ is determined by solving the equation $E_{\rho=0}[\varphi_5(\widehat{\rho}_{\mathrm{BVRSS},m})] = \alpha$. Since the value of $c(n,\alpha)$ cannot be determined analytically, we compute the values of $c(n,\alpha)$ for $n = 8, 9, 12, 16, 18, 20, 24, 28, 32$ and $\alpha = 0.01, 0.05, 0.1$ by a simulation study for $N_2(0,0,1,1,\rho=0)$, similar to Section 3. These values are given in Table 4.

Let $RV(mk^2)$ denote the ratio of the variance of $\widehat{\rho}_{\mathrm{BVSRS},m}$ and the variance of $\widehat{\rho}_{\mathrm{BVRSS},m}$. Similar to the formula (13) of Al-Saleh and Samawi (2005), for $\rho = 0$ we have

$$\begin{aligned} RV(mk^2) &= \frac{\mathrm{var}(\widehat{\rho}_{\mathrm{BVSRS},m})}{\mathrm{var}(\widehat{\rho}_{\mathrm{BVRSS},m})} \\ &= \frac{1}{1-\dfrac{1}{m}\displaystyle\sum_{s=1}^{m}\left(\left(\frac{1}{k}\sum_{i=1}^{k}\mu^2_{(i)s}\right)\left(\frac{1}{k}\sum_{j=1}^{k}\mu^2_{(j)s}\right)\right)} \end{aligned}$$

where $\mu_{(i)s}$ is the mean of ith order statistic of a sample of size k from a standard normal population in the sth repetition. By (12), we have the following approximation

$$\begin{aligned} \alpha &\approx P_{\rho=0}\left(\frac{\sqrt{m}|\widehat{\rho}_{\mathrm{BVRSS},m}|}{\sqrt{\mathrm{var}(\widehat{\rho}_{\mathrm{BVRSS},m})}} > z_{\frac{\alpha}{2}}\right) \\ &= P_{\rho=0}\left(|\widehat{\rho}_{\mathrm{BVRSS},m}| > z_{\frac{\alpha}{2}}\sqrt{\frac{\mathrm{var}(\widehat{\rho}_{\mathrm{BVSRS},m})}{mRV(mk^2)}}\right), \end{aligned}$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentile of the standard normal distribution. If we use the approximation $\mathrm{var}(\widehat{\rho}_{\mathrm{BVSRS},m}) = \frac{1}{mk^2}$, then

$$\alpha \approx P_{\rho=0}\left(|\widehat{\rho}_{\mathrm{BVRSS},m}| > z_{\frac{\alpha}{2}}\sqrt{\frac{1}{m^2k^2RV(mk^2)}}\right). \tag{14}$$

Note that $c(n,\alpha)$ of the test function $\varphi_5(\widehat{\rho}_{\mathrm{BVRSS},m})$ for testing (1) is defined by

$$\begin{aligned}\alpha &= P_{\rho=0}\left(\frac{|\widehat{\rho}_{\text{BVRSS},m}|\sqrt{mk^2-2}}{\sqrt{1-\widehat{\rho}^2_{\text{BVRSS},m}}} > c(n,\alpha)\right)\\ &= P_{\rho=0}\left(|\widehat{\rho}_{\text{BVRSS},m}| > \sqrt{\frac{1}{\frac{mk^2-2}{c^2(n,\alpha)}+1}}\right).\end{aligned} \tag{15}$$

Equating (14) and (15), we note that $c(n,\alpha)$ equals approximately to $\widetilde{c}(n,\alpha)$, where

$$\widetilde{c}(n,\alpha) = \sqrt{\frac{mk^2-2}{\frac{m^2k^2RV(mk^2)}{z^2_{\frac{\alpha}{2}}}-1}}. \tag{16}$$

Note that using table of the relevant $\mu_{(i)}$ (provided in, for example, Harter, 1961), we could calculate the quantities

$$\frac{1}{m}\sum_{s=1}^{m}\left(\left(\frac{1}{k}\sum_{i=1}^{k}\mu^2_{(i)s}\right)\left(\frac{1}{k}\sum_{j=1}^{k}\mu^2_{(j)s}\right)\right).$$

Hence $\widetilde{c}(n,\alpha)$ could be computed for each n. Thus, for a large-sample size a repeated BVRSS-based test is given by

$$\varphi_6(\widehat{\rho}_{\text{BVRSS},m}) = \begin{cases} 1 & \dfrac{\widehat{\rho}_{\text{BVRSS},m}\sqrt{mk^2-2}}{\sqrt{1-\widehat{\rho}^2_{\text{BVRSS},m}}} > \tilde{c}(n,\alpha) \\ 0 & \text{o.w} \end{cases}. \tag{17}$$

In Table 4, the values of $\widetilde{c}(n,\alpha)$ and $c(n,\alpha)$ for $n = 8, 9, 12, 16, 18, 20, 24, 28, 32$ and $\alpha = 0.01, 0.05, 0.1$ is provided, respectively. From this table we observe that the values of $\widetilde{c}(n,\alpha)$ and $c(n,\alpha)$ for $n \geq 20$ are markedly close to each other.

5 Approximation of p-value by Bootstrap Method

In this section, we approximate the p-value of the test (7) using the bootstrap method that proposed by Modarres et al. (2006) based on a BVRSS and BVSRS of size $n = k^2$ from a BVN distribution with correlation coefficient ρ. The procedure is as follows.

Table 4: The values of $\widetilde{c}(n,\alpha)$ and $c(n,\alpha)$ for $n = 8, 9, 12, 16, 18, 20, 24, 28, 32$ and $\alpha = 0.01, 0.05, 0.1$

			$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.1$	
n	m	k	$c(n,\alpha)$	$\widetilde{c}(n,\alpha)$	$c(n,\alpha)$	$\widetilde{c}(n,\alpha)$	$c(n,\alpha)$	$\widetilde{c}(n,\alpha)$
8	2	2	2.5432	1.8879	1.5139	1.2849	1.1250	1.0370
9	1	3	4.9170	3.0407	2.4901	1.8548	2.0074	1.4546
12	3	2	1.6142	1.4089	1.1845	1.0299	0.8921	0.8512
16	4	2	1.2133	1.1993	0.9211	0.8935	0.7702	0.7436
18	2	3	1.6402	1.6291	1.1758	1.1985	1.0391	0.9928
20	5	2	1.0943	1.0683	0.8139	0.8023	0.6840	0.6699
24	6	2	0.9811	0.9748	0.7410	0.7552	0.6394	0.6148
28	7	2	0.8917	0.9032	0.6945	0.6828	0.5823	0.5716
32	8	2	0.8198	0.8458	0.6514	0.6404	0.5239	0.5364
36	4	3	1.1201	1.1198	0.8399	0.8456	0.7188	0.7074
48	3	4	1.2174	1.2111	0.9167	0.9154	0.7624	0.7661
75	3	5	1.1433	1.1386	0.8613	0.8631	0.7208	0.7133
100	4	5	0.9851	0.9856	0.7450	0.7484	0.6249	0.6275

5.1 Approximation of p-value Based on BVSRS

Suppose that (X, Y) denote a population with distribution function $F_{X,Y}$, $E(X) = \mu_x$, $E(Y) = \mu_y$, $V(X) = \sigma_x^2$, $V(Y) = \sigma_y^2$ and $\mathrm{Corr}(X, Y) = \rho$. Assume that μ_x, μ_y, σ_x^2, σ_y^2 and ρ are unknown. For testing (7) based on bootstrap method, do the following steps:

1. Randomly draw k^2 elements $(X_{ij}, Y_{ij}) \sim F_{X,Y}$ and assign to each element a probability of $\frac{1}{k^2}$.

2. Select a sample of size $n = k^2$ from sample of step 1 based on BVSRS method.

3. Compute $t_{\text{BVSRS}} = \dfrac{r_{\text{BVSRS}}\sqrt{n-2}}{\sqrt{1 - r^2_{\text{BVSRS}}}}$ and call it t^*_{BVSRS}

4. Repeat step 2, N times and compute t_{BVSRS} for each repetition.

5. Compute bootstrap p-value by the following formula

$$\begin{aligned}\text{bootstrap} \quad p-\text{value} &= P(t > t^*) \\ &= \frac{\text{number of samples which satisfy } t > t^*}{N}\end{aligned}$$

6. Repeat the above steps $B = 2500$ times and approximate the p-values by the mean of the bootstrap p-values.

5.2 Approximation of p-value Based on BVRSS

Suppose that (X, Y) denote a population with distribution function $F_{X,Y}$, $E(X) = \mu_x$, $E(Y) = \mu_y$, $V(X) = \sigma_x^2$, $V(Y) = \sigma_y^2$ and $\text{Corr}(X, Y) = \rho$. Assume that μ_x, μ_y, σ_x^2, σ_y^2 and ρ are unknown. In order to test (7) based on bootstrap method, do the following steps:

1. Randomly draw k^4 elements $(X_{ij}, Y_{ij}) \sim F_{X,Y}$ and assign to each element a probability of $\frac{1}{k^4}$.

2. Select a sample of size $n = k^4$ with replacement from sample of step 1, and perform a BVRSS of size k^2 from this sample.

3. Compute $t_{\text{BVRSS}} = \frac{r_{\text{BVRSS}}\sqrt{n-2}}{\sqrt{1-r^2_{\text{BVRSS}}}}$ and call it t^*_{BVRSS}

4. Repeat step 2, N times and compute t_{BVRSS} for each repetition.

5. Compute bootstrap p-value by the following formula

$$\begin{aligned}\text{bootstrap} \quad p-\text{value} &= P(t > t^*) \\ &= \frac{\text{number of samples which satisfy } t > t^*}{N}\end{aligned}$$

6. Repeat the above steps $B = 2500$ times and approximate the p-values by the mean of the bootstrap p-values.

For computing the approximated p-value, we generate a sample from $N_2(0, 1, 1.5, 4, \rho = 0)$. In Table 5, values of approximate p-value have been computed according to the above procedures for $n = 4, 8, 9, 12, 16, 18, 20, 24, 25, 28, 32$. Considering this table we could observe that for all n, p-value of BVRSS method is less than p-value of BVSRS method.

Table 5: p-values of the BVSRS and BVRSS-based tests for a correlation coefficient of bivariate normal distribution with bootstrap method

n	m	k	BVRSS	BVSRS
1	2	4	0.1591	0.3013
2	2	8	0.2015	0.3698
1	3	9	0.2025	0.3253
3	2	12	0.1511	0.2674
4	2	16	0.1380	0.3137
2	3	18	0.1727	0.3106
5	2	20	0.1923	0.3141
6	2	24	0.1799	0.2467
1	5	25	0.1819	0.2834
7	2	28	0.1484	0.2736
8	2	32	0.1614	0.2960

6 Conclusion

The use of BVRSS is limited to situations, in which ranking of a small number of units by judgment could be carried out using negligible ranking errors. When the BVRSS is applicable, then the proposed tests based on BVRSS to test correlation coefficient of BVN distribution, is more powerful than UMPU test. In addition, if the ranking of samples is carried out exactly, then the BVRSS tests based on unrepeated samples would be more powerful than BVRSS tests based on repeated samples for the BVN distribution. However, since in practice, the error in ranking of samples could occur, it is recommended to use BVRSS tests-based on repeated samples which are more powerful than usual UMPU tests based on BVSRS.

Although we use BVN distribution to compare the power of the proposed test functions to the corresponding UMPU tests, a similar test functions could be constructed based on UMPU tests and BVRSS for any bivariate distributions.

Acknowledgments

The authors are grateful to the Editor, and two anonymous referees for making helpful comments and suggestions on an earlier version of this article. Research of Nader Nematollahi was supported by the research council of Allameh Tabataba'i University.

References

Abu-Dayyeh, W. and Muttlak, H. A. (1996), Using ranked set sampling for hypothesis tests on the scale parameters of exponential and uniform distributions. Pakistan Journal of Statistics, **12**, 131-138.

Al-Saleh, M. F. and Samawi, H. M. (2005), Estimation of the correlation coefficient using bivariate ranked set sampling with application to the bivariate normal distribution. Commun. Statist. Theor. Meth., **34**, 875-889.

Al-Saleh, M. F. and Zheng, G. (2002), Estimation of bivariate characteristics using ranked set sampling. Aust. N. Z. J. Stat., **44**(2), 221-232.

Bhoj, D. S. and Ahsanullah, M. (1996), Estimation of parameters of the generalized geometric distribution using ranked set sampling. Biometrics, **52**, 685-694.

Chen, Z., Bai, Z., and Sinha, B. K. (2004), Ranked Set Sampling: Theory and applications. New York: Springer.

Dell, T. R. and Clutter, J. L. (1972), Ranked set sampling theory with order statistics background. Biometrics, **28**, 545-555.

Fei, H., Sinha, B. K., and Wu, Z. (1994), Estimation of parameters in two-parameter Weibull and extreme-value distributions using ranked set sample. J. Statist. Res., **28**, 149-161.

Harter, H. L. (1961), Expected values of normal order statistics. Biometrika, **48**, 151-165, Correction pp. 476.

Hossain, S. S. and Khan, A. (2006), Test procedures with selected ranked set sampling. Statistica, **46**(2), 161-170.

Hossain, S. S. and Muttlak, H. A. (2001), Selected ranked set sampling. Australian and New Zealand Journal of Statistics, **43**, 311-325.

Hossain, S. S. and Muttlak, H. A. (2006), Hypothesis tests on the scale parameter using median ranked set sampling. Statistica, anno LXVI 4.

Hui, T. P., Modarres, R., and Zheng, G. (2009), Pseudo maximum likelihood estimates using ranked set sampling with applications to estimating correlation. Test, **18**, 365-380.

Lam, K., Sinha, B. K., and Wu, Z. (1994), Estimation of parameters in a two-parameter exponential distribution using ranked set sample. Ann. Inst. Statist. Math., **46**, 723-736.

Lam, K., Sinha, B. K., and Wu, Z. (1995), Estimation of location and scale parameters of a logistic distribution using ranked set sample. In: Nagaraja, Sen, Morrison, (Eds), papers in Honor of Herbert A. David. 187-197.

Li, D., Sinha Bimal, K., and Perron, F. (1999), Random selection in ranked set sampling and its applications. J. Statist. Plann. Infer., **75**, 185-201.

McIntyre, G. A. (1952), A method for unbiased selective sampling, using ranked sets. Austral. J. Agricult. Res., **3**, 385-390.

Modarres, R., Hui, T. P., and Zheng, G. (2006), Resampling methods for ranked set samples. Computational Statistics & Data analysis, **51**, 1039-1050.

Muttlak, H. A. (1997), Median ranked set sampling. J. Appl. Statist. Sci., **6**(4), 245-255.

Muttlak, H. A. and Abu-Dayyeh, W. (1998), Testing some hypotheses about the normal distribution using ranked set sample: A more powerful test. Journal of Information and Optimization Sciences, **19**, 1-11.

Ni Chuiv, N., Sinha, B. K., and Wu, Z. (1994), Estimation of the location parameter of a Cauchy distribution using a ranked set sample. Technical Report, University of Maryland Baltimore County.

Patil, G. P., Sinha, A. K., and Taillie, C. (1994), Ranked set sampling for multiple characteristics. Internat. J. Ecol. Environ. Sci., **20**, 357-373.

Rahimov, I. and Muttlak, H. A. (2003), Estimation of the population mean using random selection in ranked set samples. Statist. Probab. Lett. J. **62**, 203-209.

Samawi, H. M., Ahmed, M. S., and Abu-Dayyeh, W. (1996), Estimating of population mean using extreme ranked set samples. Biom. J., **38**(5), 577-586.

Shao, J. (1999), Mathematical Statistics. New York and Berlin: Springer.

Sinha, B. K. and Purkayastha, S. (1996), On some aspects of ranked set sampling for estimation of Normal and Exponential parameters. Statist. Decisions, **14**, 223-240.

Stokes, S. L. (1980), Inference on the correlation coefficient in bivariate normal population from ranked set sampling. J. Amer. Statist. Assoc., **75**, 989-995.

Takahasi, K. (1970), Practical note on estimation of population means based on samples stratified by means of ordering. Ann. Inst. Statist. Math., **22**, 421-428.

Takahasi, K. and Wakimoto, K. (1968), On unbiased estimators of the population mean based on the sample stratified by means of ordering. Ann. Inst. Statist. Math., **21**, 249-255.

Tseng, Y. L. and Wu, S. W. (2007), Ranked-set-sample-based tests for normal and exponential means. Commun. Statist. Simul. Computat., **36**, 761-782.

Wang, M. F. and Tseng, Y. L. (2002), Ranked-set-sample based better tests for a normal mean. J. Chinese Statist. Assoc., **40**(3), 391-418.

Zheng G. and Modarres, R. (2006), A robust estimate of the correlation coefficient for bivariate normal distribution using ranked set sampling. Journal of Statistical Planning and Inference, **136**, 298-309.

Appendix

In this appendix we compare the performance of the RSS estimator of ρ given by Al-Saleh and Samawi (2005) to three other estimators of ρ. For the sake of simplicity, we consider the case of known means and variances and let $\mu_x = \mu_y = 0$ and $\sigma_x = \sigma_y = 1$. So, we compare the performance of the $\widehat{\rho}_{\text{BVRSS}}$ given by (11) to the following three estimators of ρ. The first estimator of ρ is $\widehat{\rho}_{\text{SRS}} = r_{\text{BVSRS}}$ given in (10) with $n = mk^2$. The second estimator of ρ is the MLE based on RSS given by Stokes (1980). If $X_{(i)j}$ is the ith smallest unit in the ith group $(i = 1, 2, \cdots, k^2)$ of simple random samples, and in the jth cycle $(j = 1, 2, \cdots, m)$; and $Y_{[i]j}$ is the value of Y's concomitant to $X_{(i)j}$, then the $\widehat{\rho}_{\text{MLE}}$ can be found by solving the cubic equation

$$h(\rho) = mn\rho(1-\rho^2)+(1+\rho^2)\sum_{j=1}^{m}\sum_{i=1}^{k^2} X_{(i)j}Y_{[i]j}-\rho\left[\sum_{j=1}^{m}\sum_{i=1}^{k^2} X_{(i)j}^2 \sum_{j=1}^{m}\sum_{i=1}^{k^2} Y_{[i]j}^2\right] = 0$$

The third estimator was proposed by Zheng and Modarres (2006) and defined by

$$\widehat{\rho}_{\text{RSS}} = \frac{1}{mk^2}\sum_{j=1}^{m}\sum_{i=1}^{k^2} X_{(i)j}Y_{[i]j}.$$

We generate BVRSS and BVSRS from $N_2(0, 0, 1, 1, \rho)$ with $\rho = 0.0, 0.1, 0.2, 0.3, 0.5, 0.7$ and $j = 1, 2, \cdots, 8$, $m = 1, 2$, and repeat the procedure 1000 times. In our simulation, all rankings are assumed to be perfect. The simulation results are reported in Table 6. In this table, the values E_1, E_2 and E_3 are relative efficiencies (RE) of $\widehat{\rho}_{\text{BVRSS}}$ with respect to $\widehat{\rho}_{\text{RSS}}$, $\widehat{\rho}_{\text{MLE}}$ and r_{BVSRS}, respectively. The RE is the ratio of the respective MSE's. $RE > 1$ indicates that the $\widehat{\rho}_{\text{BVRSS}}$ is more efficient than the alternative estimators. From Table 6, we see that $\widehat{\rho}_{\text{BVRSS}}$ is more efficient than $\widehat{\rho}_{\text{MLE}}$ and r_{BVSRS} for all ρ and more efficient than $\widehat{\rho}_{\text{RSS}}$ for almost all ρ.

Table 6: Relative efficiency of $\widehat{\rho}_{\text{BVRSS}}$ w.r.t $\widehat{\rho}_{\text{RSS}}(E_1)$, $\widehat{\rho}_{\text{MLE}}(E_2)$ and $r_{\text{BVSRS}}(E_3)$

			$N(0,0,1,1,0)$			$N(0,0,1,1,0.1)$		
m	k	n	E_1	E_2	E_3	E_1	E_2	E_3
1	2	4	1.0697	1.0992	1.3571	1.1357	1.1257	1.4579
2	2	8	1.1454	1.1230	1.4930	1.0775	1.1143	1.4002
1	3	9	1.2559	1.1008	1.7451	1.2139	1.1243	1.7591
3	2	12	1.0041	1.0976	1.3557	1.2024	1.1084	1.3947
4	2	16	1.0014	1.1049	1.1695	1.1097	1.1095	1.3710
2	3	18	1.0010	1.1135	1.0285	1.0944	1.1017	1.3659
5	2	20	1.0711	1.1066	1.2209	1.0749	1.0976	1.3622
6	2	24	1.2167	1.0999	1.5496	1.0389	1.0762	1.2508
7	2	28	1.0325	1.0798	1.1609	1.1185	1.0537	1.2137
8	2	32	1.0491	1.0732	1.2032	1.0954	1.0071	1.4088
			$N(0,0,1,1,0.2)$			$N(0,0,1,1,0.3)$		
m	k	n	E_1	E_2	E_3	E_1	E_2	E_3
1	2	4	1.1176	1.1185	1.4837	1.1279	1.1344	1.5231
2	2	8	1.0284	1.1253	1.4826	1.1341	1.1296	1.4259
1	3	9	1.3428	1.1139	1.7303	1.1432	1.1188	1.6872
3	2	12	1.0764	1.1008	1.4987	1.2014	1.1150	1.3186
4	2	16	1.0081	1.1010	1.4138	1.1007	1.1165	1.2905
2	3	18	1.0037	1.0894	1.0682	1.0576	1.1024	1.3334
5	2	20	1.0179	1.0953	1.4554	1.0457	1.1074	1.3962
6	2	24	1.0773	1.0986	1.2666	1.1290	1.1160	1.2822
7	2	28	1.0777	1.0840	1.2687	1.1448	1.0965	1.2129
8	2	32	1.1334	1.0770	1.2174	1.0447	1.0873	1.3436
			$N(0,0,1,1,0.5)$			$N(0,0,1,1,0.7)$		
m	k	n	E_1	E_2	E_3	E_1	E_2	E_3
1	2	4	1.0808	1.1223	1.6642	1.1266	1.1104	1.9143
2	2	8	0.9333	1.1375	1.3527	0.9574	1.0934	1.3870
1	3	9	1.2811	1.1146	1.8897	1.0049	1.0962	1.2462
3	2	12	0.9858	1.1080	1.3432	1.0003	1.0705	1.0934
4	2	16	0.9989	1.1152	1.2906	0.9875	1.0846	1.1108
2	3	18	0.9718	1.1001	1.1407	0.9703	1.0612	1.1046
5	2	20	1.0413	1.0948	1.2443	0.9822	1.0518	1.1374
6	2	24	0.9903	1.0856	1.1565	0.9809	1.0594	1.0738
7	2	28	0.9674	1.0679	1.2050	0.9653	1.0636	1.0465
8	2	32	0.9936	1.0700	1.1863	0.9641	1.0580	1.0158